

DOCUMENT RESUME

ED 411 635

EC 305 826

AUTHOR Plucker, Jonathan A.
 TITLE Student Responses to Creativity Checklists: Evidence of Content Generality.
 PUB DATE 1996-10-17
 NOTE 18p.
 PUB TYPE Reports - Research (143)
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS *Check Lists; College Freshmen; Correlation; *Creativity; Factor Analysis; High School Seniors; High Schools; Higher Education; Multivariate Analysis; Recreational Activities; Research Methodology; *Statistical Analysis; *Student Attitudes

ABSTRACT

This study re-examined three previous studies on creativity in an attempt to address the question of whether creativity is content general or content specific. Statistical re-examinations were conducted that involved inferential statistical analyses (factor analysis and multivariate analysis) of correlation matrices representing student scores on creative activity checklists. Each of the three reanalyses includes a brief description of the original study and a summary of the researchers' original conclusions followed by a re-examination of the data. The studies involved analysis of: (1) activity checklists of high school seniors who qualified to be National Merit Finalists; (2) activity checklists of college freshmen; and (3) revised checklists completed by undergraduate students. The overall re-analysis study found a number of discrepancies requiring further research, but suggests that in regard to the content general or content specific question there is a method effect. Performance assessments produce evidence of content and task specificity, while creativity checklists suggest that creativity is applied generally across disciplines. (Contains 36 references.) (DB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

Running Head: CONTENT GENERALITY OF CREATIVITY CHECKLISTS

ED 411 635

Student Responses to Creativity Checklists:

Evidence of Content Generality

Jonathan A. Plucker¹

University of Maine

October 17, 1996

5766 Shibles Hall

University of Maine

Orono, ME 04469-5766

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
- ☐ Minor changes have been made to
improve reproduction quality.

- Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

PLUCKER

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

EC 305826

Student Responses to Creativity Checklists:

Evidence of Content Generality

Abstract

The question of whether creativity is content general or content specific is one of the most controversial issues in contemporary creativity research, with recent studies providing support for both positions. Reanalysis of data from three previously published studies generally suggest evidence of content generality. The results are analyzed in light of other studies, and a hypothesis regarding the content generality-specificity of creativity is provided in order to guide future research efforts.

Student Responses to Creativity Checklists:

Evidence of Content Generality

Although issues of content generality of creativity have only recently become the subject of comprehensive research efforts, assumptions of content generality or specificity permeate research and educational efforts involving most aspects of creative production. Indeed, the psychometric study of creativity, dating from the early 1950s, is dominated by the generality perspective. A majority of divergent thinking tests (Guilford, 1967; Torrance, 1974; Wallach & Kogan, 1965) are predicated on the belief that creative thinking is manifest across content areas, with scores for various dimensions such as fluency, flexibility, and originality applying across art, writing, science, and other topic areas. For example, high originality scores on a divergent thinking test with science content are expected to generalize to high originality scores on literature or even domain-neutral divergent thinking tests. Conversely, many researchers, using both psychometric and other methodologies, study creativity under the assumption of discipline specificity (see Runco, 1987).

In the last 10 to 15 years, however, the content generality of creativity steadily gained the attention of researchers and became perhaps the most controversial topic in the study of creativity. Baer (1993a, 1993b, 1994d), based on his research with creative products and the Consensual Assessment Technique (CAT; Amabile, 1983, 1996), strongly criticizes the use of divergent thinking tests due to their theoretical basis of generality. Baer's research (1994a, 1994b, 1994c) provides evidence that creative production is not only content specific but also task specific within content areas. This finding obviously stands in sharp contrast to the work of many divergent thinking proponents and is the subject of occasional criticism (Cramond, 1994; Kogan, 1994), but other research also supports the notion of content specificity (Holland, 1961; Holland & Astin, 1962; Runco, 1987, 1989).

After considering this research, Plucker and Renzulli (in press) concluded that questions regarding the general or domain specific nature of creativity have yet to be answered satisfactorily. Researchers involved in the generality-specificity debate employ diverse

methodologies that have specific strengths and limitations, and many of the studies are exploratory in nature. For example, many traditional psychometric studies do not utilize multivariate statistical analyses, and performance-based assessments (e.g., the CAT) generally show evidence of task specificity regardless of the process skills that are the focus of the assessment (Dunbar, Koretz, & Hoover, 1991; Linn & Burton, 1994).

Determining whether creativity is manifest only in specific content areas or to specific tasks as opposed to most areas of productivity holds important implications for education and talent development as well as the empirical study of creativity. If evidence of task specificity were found among children, creativity training programs could focus on specific areas of creative accomplishment in which students hold particular promise. If creativity is domain general, educators could focus upon general creative process skills in their efforts to increase children's capacity for creative productivity. In a similar vein, further evidence of task specificity may explain the perceived lack of predictive and discriminant validity for content-neutral divergent thinking tests. In order to inform practice and future research efforts, increased attention should be focused on issues of creativity specialization or lack thereof.

Purpose

Many creativity studies published prior to the 1980s, especially those involving administrations of creative activity or behavior checklists, contain few instances of multivariate statistical analysis. Most of these studies only include bivariate analyses, providing sufficient vagueness to allow a wide range of conclusions to be drawn. Given the rapid progression of statistical analysis and technological advances in related computer programs over the past few decades, this limitation of previous studies is understandable and correctable. Plucker and Renzulli (in press) recently recommended that statistically elegant techniques be applied to the reanalysis of previously published correlation matrices derived from scores on creativity instruments. The purpose of this paper is to report the results of several statistical re-examinations of previously published (and underanalyzed) data. In each of the reanalyzed studies, correlation matrices representing student scores on creative activity checklists were

previously interpreted without additional inferential statistical analyses (e.g., factor analysis). Multivariate analysis of the three data sets will allow more direct responses to questions of creative content-generality.

Method

Studies were selected for reanalysis based on two criteria. First, each study needed to utilize similar methodology -- in this case, administration of creative activity inventories. The research of Holland and Nichols (1964) and Holland and Richards (1965) used a creative activity checklist developed by Holland (1961; Holland & Astin, 1962), and Hocevar (1976) employed his revision (Hocevar, 1979) of the Holland instrument. Researchers and reviewers frequently conclude that attainment checklists exhibit relatively high levels of reliability and validity evidence when used to measure creativity (Bull & Davis, 1980; Davis, 1975; Hocevar, 1981; Hocevar & Bachelor, 1989; Milgram & Milgram, 1976; Wallach, 1976). Second, each study had to be frequently cited in the creativity literature. The work of Hocevar and of Holland and his colleagues forms the foundation for most creative activity checklist studies conducted in recent decades, thereby satisfying the second criterion.

Because only correlation matrices are available in the three articles (other descriptive statistics such as means and standard deviations are not provided), exploratory factor analysis was used to reanalyze the data (Pedhazur & Schmelkin, 1991; Tabachnick & Fidell, 1989). Each of the following reanalyses includes a brief description of the original study and summary of the researchers' original conclusions followed by a re-examination of the data. Since the correlation matrices analyzed in this paper are available in the original articles, the matrices are not included here.

Results

Reanalysis 1

Holland and Nichols (1964) administered their activity checklists to high school seniors who qualified to be National Merit Finalists. The checklists included questions regarding creative achievement in science, art, leadership, drama, writing, and music. Reliability

estimates (KR-20) ranged from .48 to .75 for boys and .58 to .86 for girls. Correlation matrices for the checklist scale scores were constructed separately for each gender and in both cases included relatively small correlations among students' scores on the six scales. Holland and Nichols concluded that "the criteria are relatively independent of one another with the following exceptions: for boys -- leadership, dramatics, and writing achievements are moderately correlated with one another; for girls -- art, writing, music, and leadership tend to be intercorrelated" (p. 58). These moderate correlations ranged from .21 to .30, with the remaining correlations ranging from -.12 to .19.

Correlation matrices for boys and girls were factor analyzed separately using a variety of extraction techniques (i.e., maximum likelihood, principal axis, and alpha factoring) with oblimin rotation. Since results were very similar across extraction techniques, results are reported for maximum likelihood extraction. For the boys, three factors were extracted based on the Kaiser normalization and scree test. Pre-rotation scale score communalities ranged from .18 to .39, factors accounted for little pre-rotation variance (17.0%, 7.2%, 5.3%). Post-rotation structure loadings and factor correlations are included in Table 1. The first factor is characterized by high loadings for dram, leadership, and writing scores, while the second factor appears to most highly correlate with music scores. While both science and art scores load on the third factor, art has nearly uniform loadings across all three factors.

For girls, two factors were extracted, communalities ranged from .08 to .99, and factors again accounted for little pre-rotation variance (18.7%, 13.0%). Loadings and factor correlations for girls are also included in Table 1. In contrast to the results for boys, scores on the art scale load very highly on the first factor. Leadership scores have the highest loading on the second factor, with all other scale scores correlating only moderately with the second factor. Both for boys' and for girls' scores, factor correlations are low.

Reanalysis 2

Holland and Richards (1965) administered the same activity checklists (the writing scale was renamed the literary scale) to college freshmen at 24 college campuses. These

students had taken the American College Test battery during their senior year of high school. Reliability estimates (KR-20) ranged from .72 to .84 for males and .65 to .81 for females. Correlations among the scale scores ranged from .18 to .50, with 87% of the correlations in excess of .25. Holland and Richards concluded that "there is some generality to artistic, scientific, and social accomplishments" (p. 170).

During factor analyses of the 1965 data, the three extraction techniques again provided similar results, and those produced with maximum likelihood extraction will be reported here. For males, the scree test indicated the presence of one factor, with initial eigenvalues of 2.88, .84, .70, .61, .52, and .43, respectively, for the six potential factors. The factor accounted for 38.0% of the variance, and communalities ranged from .30 (Music) to .53 (Drama). For females, eigenvalues for the six potential factors (2.57, .89, .80, .66, .57, .50) suggested a one factor solution. The factor accounted for 32.1% of the variance with communalities ranging from .18 (Art) to .54 (Drama). In both sets of analyses, two and three factor solutions (when estimable) were characterized by high factor correlations and high loadings of all variables upon all factors.

However, the Holland and Richards study (1965) included information that was not included in the Holland and Nichols (1964) investigation: specific reliability estimates for each scale. Using this added data, the correlations in the Holland and Richards (1965) article were corrected for attenuation due to measurement error (Nunnally, 1970). Analysis of the corrected correlation matrices produced one factor solutions (accounting for 49.3% of the variance for males and 44.1% for females). As was generally the case in the analyses of the attenuated data, the Art scale scores had the smallest communalities (males, .34; females, .20) and Drama scores had the largest communalities (males, .74; females, .76). Loadings for the four solutions are presented in Table 2.

Reanalysis 3

Hocevar (1976) revised the Holland checklists by adding items and reorganizing the corresponding scales (see Hocevar, 1979, for a detailed description of the revision and a copy

of the instrument). The resulting Creative Behavior Checklist includes scales of creative achievement in the fine arts, crafts, performing arts, math-science, literature, and music. Undergraduate students completed the checklists in introductory psychology and educational psychology classes, with reliability estimates ranging from .63 to .90. Hocevar corrected the scale score correlations for attenuation due to measurement error, resulting in correlations ranging from .17 to .76, with 73% between .30 and .68 in magnitude. He noted that the results suggest "a generalized disposition to distribute one's creative efforts across areas" (p. 870).

As was the case with the Holland and Richards (1965) data, a one factor solution appears most viable for the Hocevar (1976) data regardless of the technique used to extract the factors (e.g., initial eigenvalues of 2.88, .91, .85, .66, .47, and .23 for boys and 3.52, .70, .66, .51, .40, and .20 for girls with maximum likelihood extraction). Variable loadings, factor eigenvalues, and variance accounted for by the factor are included in Table 3.

Discussion

The discrepancy between the reanalysis of the Holland and Nichols (1964) data and the reanalyses of the Holland and Richards (1965) and Hocevar (1976) data may be due to several factors. First, the homogeneity of the sample with respect to ability is often directly related to the psychometric quality of creativity measurements (e.g., Runco, 1985). Since the 1964 study included high achieving high school students and the latter two investigations studied a general ability sample of undergraduates (i.e., more heterogeneous samples), this explanation appears probable. A related factor is the possible influence of response bias in the 1964 study, which (in contrast to the two later studies) relied upon surveys distributed through the mail. Given the relatively low response rate (68% vs. approximately 100% in 1965 and 1976), the conclusion that students with creative specialties were predisposed to respond while generally creative students were not motivated to return surveys is plausible. Without a detailed inspection of means and standard deviations, gathering further evidence of response bias is not possible.

A third possibility is that psychometric quality of the checklists increased with each study, producing more unitary and less ambiguous results with each successive administration. Correction for attenuation due to measurement error clearly improved the one factor solution for the Holland and Richards (1965) data, providing support for the psychometric improvement hypothesis. Collectively, the possible explanations suggest that the discrepant results were related to differences in the psychometric quality of the creative checklists and/or response bias and not ability differences in group composition.

Implications for Research on Creativity and Domain Generality

Work with creative checklists needs to be replicated both with diverse and homogeneous samples of students and adults. In addition, investigation of content generality using other creativity measures (e.g., divergent thinking tests, personality inventories, product rating scales) will further elucidate issues of content generality/specificity. Furthermore, gender differences in the reanalyses of the Holland and Richards and Hocevar data are inconclusive, and replication of this study will allow gender comparisons of factor structure utilizing confirmatory factor analysis.

However, given the existing literature in this area, a preliminary hypothesis regarding the generality of creativity across disciplines can be proposed. A method effect is clearly present in the empirical creativity literature: Performance assessments produce evidence of content and task specificity, while creativity checklists suggest that creativity is applied generally across disciplines. A case in point is the study conducted by Runco (1987)², in which creativity checklist responses were compared to quality ratings of creativity. The quality ratings were scored using a technique not unlike the CAT, and while the checklist responses suggested content generality, the quality ratings (based on alternative assessments) implied the existence of content specificity.

These results are seemingly in conflict, but they may also be providing creativity researchers with a glimpse of a social science application of the Heisenberg uncertainty principle that is well-known to physics students. Heisenberg posited that merely attempting to

measure a subatomic particle changed its properties. In the same vein, the methods we use to measure creativity -- and the conditions under which these measurements are attempted -- may effectively predetermine the results with regard to content/task generality.

This hypothesis may appear to be a not-so-clever way of stating that researchers will never elucidate the role of task or domain specificity in creativity. However, a similar position in the area of human intelligence does not restrict investigation into the plausibility of 'g' and, rather, serves as a focal point of research efforts. If the hypothesis presented in this paper is supported by future research efforts, creativity researchers should be less concerned with proving or disproving that creativity is manifest generally across domains and tasks. Instead, much as Perkins and Salomon (1989) recommend for cognitive skills in general, creativity researchers should attempt to determine under which measurement, environmental, and age-related conditions creativity is domain and task specific and under which conditions creativity is generally applied.

Author Notes

¹The author thanks Dennis Hocevar for suggestions made regarding corrections for attenuation and Mark Runco for discussing his use of creative activity checklists. However, the analyses and resulting interpretations are solely those of the author.

²Surprisingly, Runco's study is one of very few examples of combined traditional/alternative assessment strategies in creativity research. Most of the other studies, such as that reported by Skager, Schultz, and Klein (1965), did not examine issues of content or task specificity.

References

- Amabile, T. M. (1983). The social psychology of creativity. New York: SpringerVerlag.
- Amabile, T. M. (1996). Creativity in context: Update to the social psychology of creativity. Boulder, CO: Westview.
- Baer, J. (1993a). Divergent thinking and creativity: A task-specific approach. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baer, J. (1993b, December/January). Why you shouldn't trust creativity tests. Educational Leadership, 80-83.
- Baer, J. (1994a). Divergent thinking is not a general trait: A multi-domain training experiment. Creativity Research Journal, 7, 35-36.
- Baer, J. (1994b). Generality of creativity across performance domains: A replication. Perceptual and Motor Skills, 79, 1217-1218.
- Baer, J. (1994c). Performance assessments of creativity: Do they have long-term stability? Roeper Review, 17, 7-11.
- Baer, J. (1994d, October). Why you still shouldn't trust creativity tests. Educational Leadership, 72-73.
- Bull, K. S., & Davis, G. A. (1980). Evaluating creative potential using the Statement of Past Creative Activities. Journal of Creative Behavior, 14, 249-257.
- Cramond, B. (1994, October). We can trust creativity tests. Educational Leadership, 70-71.
- Davis, G. A. (1975). In frumious pursuit of the creative person. Journal of Creative Behavior, 9, 75-87.
- Dunbar, S. G., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessment. Applied Measurement in Education, 4, 289-303.
- Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Hocevar, D. (1976). Dimensionality of creativity. Psychological Reports, 39, 869-870.

Hocevar, D. (1979, April). The development of the Creative Behavior Inventory. Paper presented at the annual meeting of the Rocky Mountain Psychological Association. (ERIC Document Reproduction Service No. ED 170 350)

Hocevar, D. (1981). Measurement of creativity: Review and critique. Journal of Personality Assessment, 45, 450-464.

Hocevar, D., & Bachelor, P. (1989). A taxonomy and critique of measurements used in the study of creativity. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), Handbook of creativity (pp. 53-75). New York: Plenum Press.

Holland, J. L. (1961). Creative and academic performance among talented adolescents. Journal of Educational Psychology, 52, 136-147.

Holland, J. L., & Astin, A. W. (1962). The prediction of the academic, artistic, scientific, and social achievement of undergraduates of superior scholastic aptitude. Journal of Educational Psychology, 56, 165-174.

Holland, J. L., & Nichols, R. C. (1964). Prediction of academic and extracurricular achievement in college. Journal of Educational Psychology, 56, 165-174.

Holland, J. L., & Richards, J. M., Jr. (1965). Academic and nonacademic accomplishment: Correlated or uncorrelated? Journal of Educational Psychology, 56, 165-174.

Kogan, N. (1994). Diverging from divergent thinking [book review]. Contemporary Psychology, 39, 291-292.

Linn, R. L., & Burton, E. (1994). Performance-based assessment: Implications of task specificity. Educational Measurement: Issues and Practices, 13(1), 5-8, 15.

Milgram, R. M., & Milgram, N. A. (1976). Creative thinking and creative performance in Israeli students. Journal of Educational Psychology, 68, 255-259.

Nunnally, J. C., Jr. (1970). Introduction to psychological measurement. New York: McGraw-Hill.

Pedhazur, E. J., & Schmelkin, L. P. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Lawrence Erlbaum.

Perkins, D. N., & Salomon, G. (1989, January/February). Are cognitive skills context-bound? Educational Researcher, 16-25.

Plucker, J. A., & Renzulli, J. S. (in press). Psychometric approaches to the study of human creativity. In R. J. Sternberg (Ed.), Handbook of human creativity. New York: Cambridge University Press.

Runco, M. A. (1985). Reliability and convergent validity of ideational flexibility as a function of academic achievement. Perceptual and Motor Skills, 61, 1075-1081.

Runco, M. A. (1987). The generality of creative performance in gifted and nongifted children. Gifted Child Quarterly, 31, 121-125.

Runco, M. A. (1989). The creativity of children's art. Child Study Journal, 19, 177-189.

Skager, R. W., Schultz, C. B., & Klein, S. P. (1965). Quality and quantity of accomplishments as measures of creativity. Journal of Educational Psychology, 56, 31-39.

Tabachnick, B. G., & Fidell, L. S. (1989). Using multivariate statistics (2nd ed.). New York: HarperCollins.

Torrance, E. P. (1974). Torrance Tests of Creative Thinking: Norms-technical manual. Lexington, MA: Ginn and Company.

Wallach, M. A. (1976, January-February). Tests tell us little about talent. American Scientist, 57-63.

Wallach, M. A., & Kogan, N. (1965). Modes of thinking in young children: A study of the creativity-intelligence distinction. New York: Holt, Rinehart and Winston.

Table 1

Structure Matrix Loadings and Factor Correlations for Holland & Nichols (1964) Data

Scale	Boys ^a			Scale	Girls ^b	
	Factor 1	Factor 2	Factor 3		Factor 1	Factor 2
Drama	.56	.06	-.06	Art	.99	.33
Leader.	.55	.19	.13	Leader.	.05	.65
Writing	.55	.09	.27	Writing	.20	.37
Music	.08	.62	.05	Drama	.22	.37
Science	.05	.04	.48	Music	.03	.30
Art	.25	.29	.30	Science	.15	.28

Note. Maximum likelihood extraction with oblimin rotation.

^a n=239

^b n=259

Table 2

Factor Loadings for Reanalysis of Holland & Richards (1965) Data

Scale	Boys ^a		Girls ^b	
	Attenuated	Corrected	Attenuated	Corrected
Drama	.73	.86	.74	.87
Leadership	.57	.70	.60	.76
Literature	.69	.80	.61	.73
Science	.58	.63	.50	.54
Music	.55	.59	.47	.53
Art	.56	.59	.43	.45
Eigenvalue	2.88	3.44	2.57	3.15
% Variance Accounted For	38.0	49.3	32.1	44.1

^a n=3770^b n=3492

Table 3

Factor Loadings for Reanalysis of Hocevar (1976) Data

Scale	Boys ^a	Girls ^b
Fine Arts	.81	.89
Literature	.78	.73
Crafts	.73	.81
Performing Arts	.53	.64
Science-Math	.41	.58
Music	.31	.55
Eigenvalue	2.88	3.52
% Variance Accounted For	39.2	50.6

^a n=110^b n=129